



Media Integration and Communication Center

Università degli Studi di Firenze

Il suono

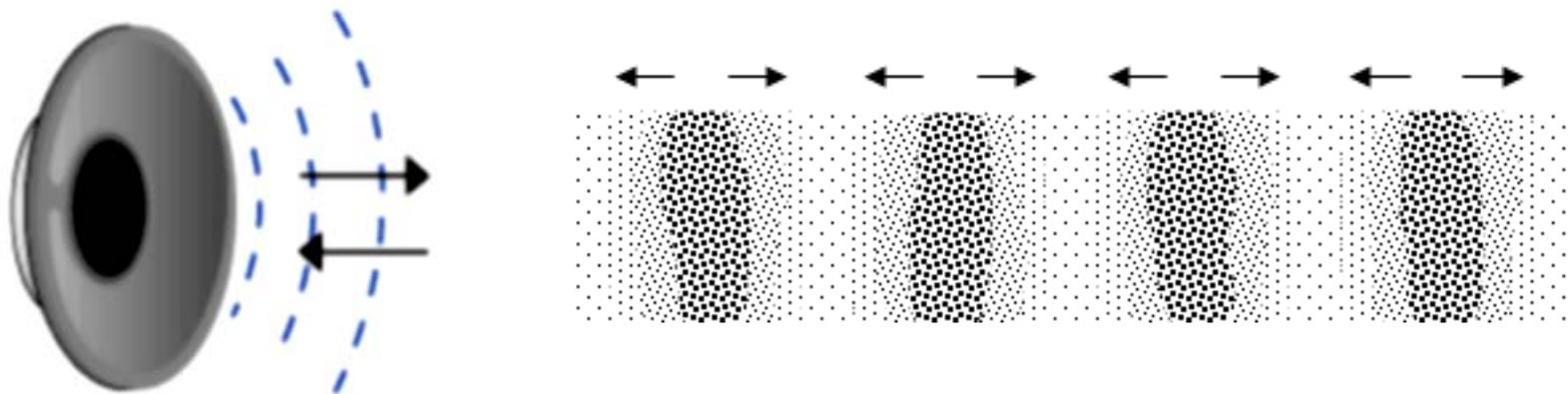
**PROGETTAZIONE E
PRODUZIONE MULTIMEDIALE**

**Prof. Alberto Del Bimbo
Dipartimento Sistemi e Informatica
Università degli Studi di Firenze**

Gianpaolo D'Amico

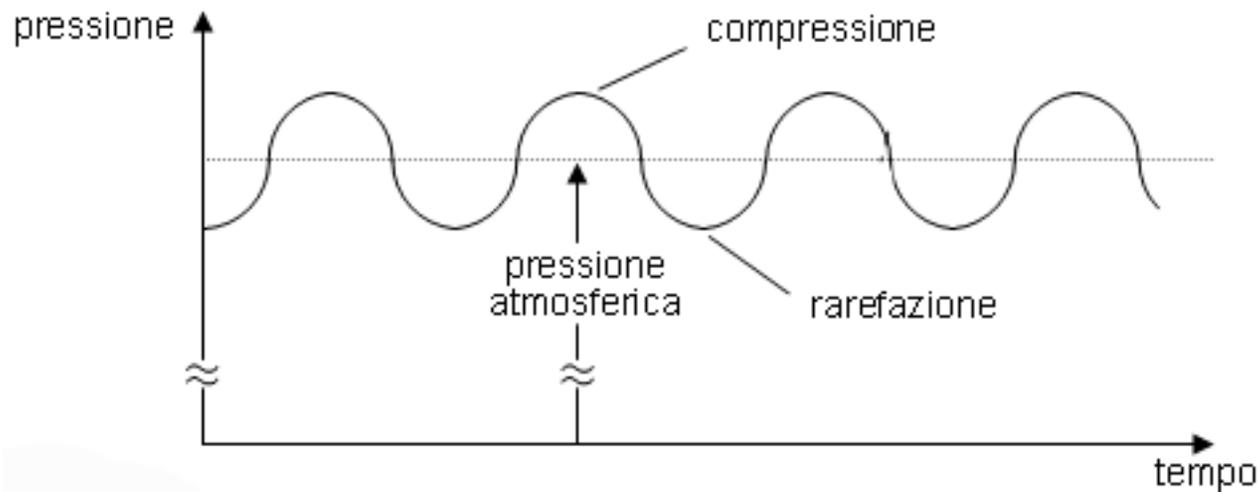
Il suono

- il suono è un fenomeno fisico prodotto da vibrazioni della materia
- come la materia vibra si creano variazioni della pressione atmosferica
- il suono non può propagarsi nel vuoto
- la percezione del suono è legata alla vibrazione delle strutture interne del nostro orecchio, causata da variazioni di pressione atmosferica.



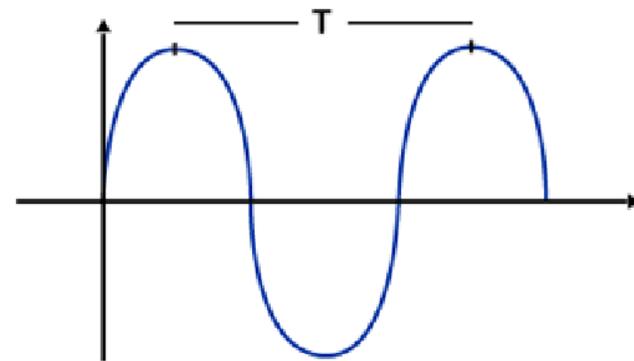
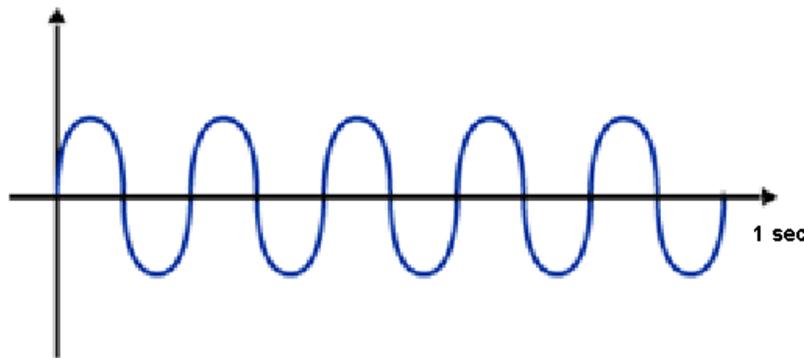
Il suono

- Le variazioni (o fluttuazioni) di pressione possono essere grandi o piccole e possono aver luogo rapidamente o lentamente.
- In ogni caso la natura della variazione di pressione e' ciclica, nel senso che l'andamento del valore di pressione si ripete con valori analoghi nel tempo.



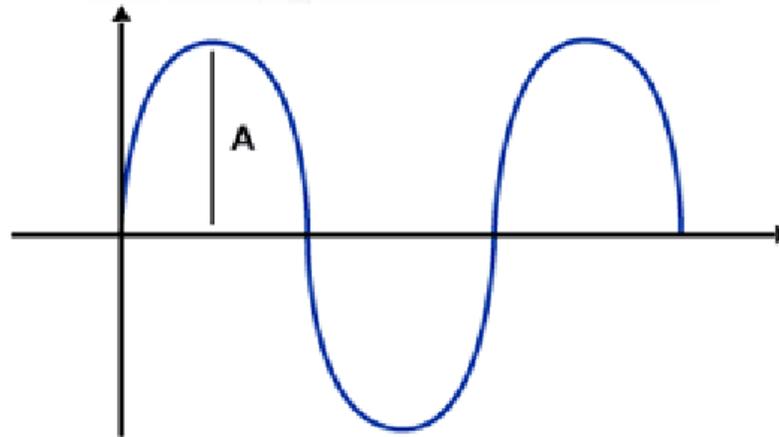
Frequenza o altezza

- La velocità con cui i valori di pressione fluttuano ciclicamente viene espressa dalla **FREQUENZA** della vibrazione: è il numero di cicli presenti in un secondo
- La durata nel tempo di ogni ciclo della oscillazione viene detta **PERIODO**.
- Frequenza (f) e periodo (T) sono legati dalla relazione $f=1/T$. Tipicamente, la frequenza si misura in Hertz (Hz), equivalendo un Hertz ad una frequenza di un ciclo per secondo.



Ampiezza o intensità

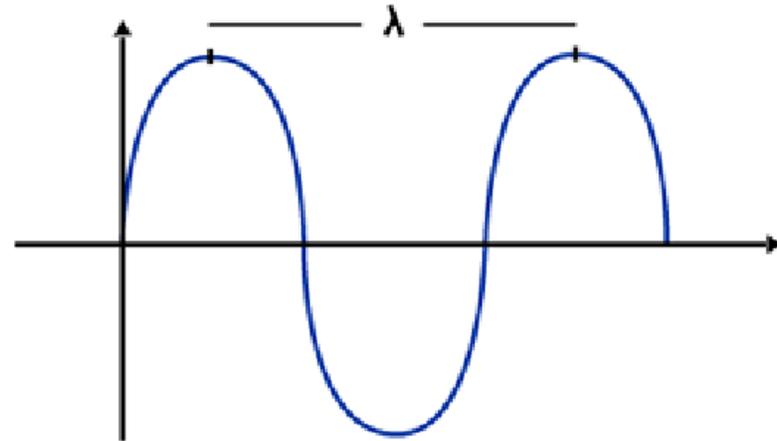
- L'altro valore che caratterizza la fluttuazione di pressione e' la sua **AMPIEZZA**, che è la misura dello scostamento massimo dalla posizione di equilibrio



Lunghezza d'onda

- E' la distanza tra due punti equivalenti lungo l'oscillazione:

$$\lambda = \frac{C}{f}$$



Dove: **C** è la velocità del suono nel mezzo, **f** la frequenza

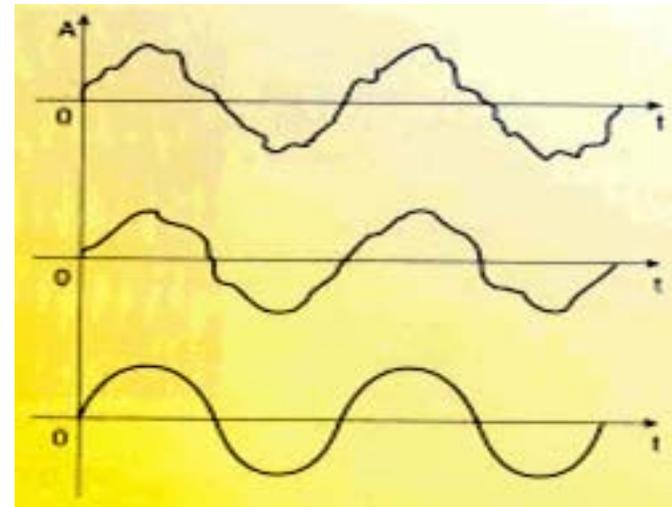
Timbro

- "elemento caratteristico di un suono che consente di differenziarlo da altri suoni aventi la stessa frequenza e intensità. E' grazie al timbro che può essere riconosciuta una voce dall'altra, oppure una stessa nota musicale emessa da due strumenti diversi" (L.Barone,G.Fanelli, A.M.Franco, 1996, p.54, *L'integrazione scolastica e sociale dei bambini minorati dell'udito*, UTET, Torino1996.).

- Il timbro rappresenta la qualità del suono e dipende essenzialmente dalla forma d'onda dello stesso e dalla natura del corpo che lo emette

- Permette di distinguere suoni emessi da sorgenti diverse, anche se essi hanno la stessa frequenza e la stessa intensità'.

- Ciascun strumento musicale ha un timbro diverso.



Frequenza e ampiezza (1)

- Il valore 0 indicato sull'asse della pressione corrisponde al valore medio della pressione dell'aria.
- Il grafico dell'andamento dice quindi che il valore della pressione aumenta gradualmente sino al tempo 1.25 msec in cui raggiunge il valore massimo di 100.
- Quindi decresce sino al tempo 3.75 msec in cui raggiunge il valore minimo -100 e ritorna a zero al tempo 5.0 msec prima di iniziare il secondo ciclo.



Frequenza e ampiezza (2)

- Ampiezza e frequenza della forma d'onda hanno effetto sul suono percepito
 - Variazioni di piccola ampiezza producono suoni di bassa intensità: il **volume** del suono cresce col crescere della sua ampiezza.
 - La frequenza della forma d'onda è invece legata alla tonalità del suono percepito: al crescere della frequenza cresce il **tono**.

Proprietà fisiche e caratteristiche percettive 1

- Le proprietà fisiche di ampiezza e frequenza corrispondono alle caratteristiche percettive di **volume** e **tono**.
- Tuttavia il legame tra queste grandezze non è costante.
- Da un punto di vista percettivo la nostra capacità di percepire suoni è limitata a quelli di frequenza compresa tra **20 Hz** e **22.000 Hz** circa.

Proprietà fisiche e caratteristiche percettive 1

- Nei limiti di quelli che sono i suoni da noi percepiti, **la relazione tra proprietà' percepite e proprietà' fisiche non e' una relazione lineare**
- Aumentando l'ampiezza di una forma d'onda di una uguale grandezza, non si ottengono uguali incrementi di volume (il volume sembra aumentare di meno via via che diventa piu'elevato).
- Analoga caratteristica vale per la frequenza: ad aumenti uguali di frequenza non corrispondono uguali incrementi di tono (l'incremento di tono sembra via via piu'piccolo col crescere della frequenza).

Proprietà fisiche e caratteristiche percettive 2

- Esempio: si noti come, sebbene gli incrementi di frequenza siano costanti (25 Hz), via via che la frequenza cresce, gli incrementi di tono sembrano diminuire.

200Hz 225Hz 250Hz 275Hz 300Hz
325Hz 350Hz 375Hz 400Hz 425Hz
450Hz 475Hz

Proprietà fisiche e caratteristiche percettive 3

- Di solito vengono quindi usate delle scale di rappresentazione tali da riflettere uguali differenze percettive.
 - Per la frequenza, una di queste scale e' la scala di **Mel**. Uguali differenze sulla scala di Mel corrispondono ad uguali differenze percepite di tono, ma non uguali differenze di frequenza.

Proprietà fisiche e caratteristiche percettive 3

- Per il volume, la scala di rappresentazione impiegata e' in **Decibel** (dB). In questa scala le ampiezze sono rappresentate in scala logaritmica:

$$L_1 = 10 \cdot \log \left(\frac{W_1}{W_2} \right) dB$$

dove: W_1 e W_2 sono potenze sonore espresse in Watt,
con W_2 potenza di riferimento;

- Nella scala in decibel, 0.0 dB corrisponde ad un suono alla soglia dell'udibilita', mentre 130 dB ad un suono alla soglia del dolore.
- Incrementi di 1 dB corrispondono a *Just Noticeable Differences* di volume (definite come le piu'piccole differenze di volume che vengono percepite in media il 50% delle volte).

Proprietà fisiche e caratteristiche percettive 4

- La scala logaritmica si differenzia dalla scala lineare per il fatto che la proporzionalità tra le due grandezze non è costante ma ha un andamento appunto logaritmico.
- La tabella seguente evidenzia la diversa corrispondenza tra due grandezze X e Y legate da una relazione lineare e logaritmica:

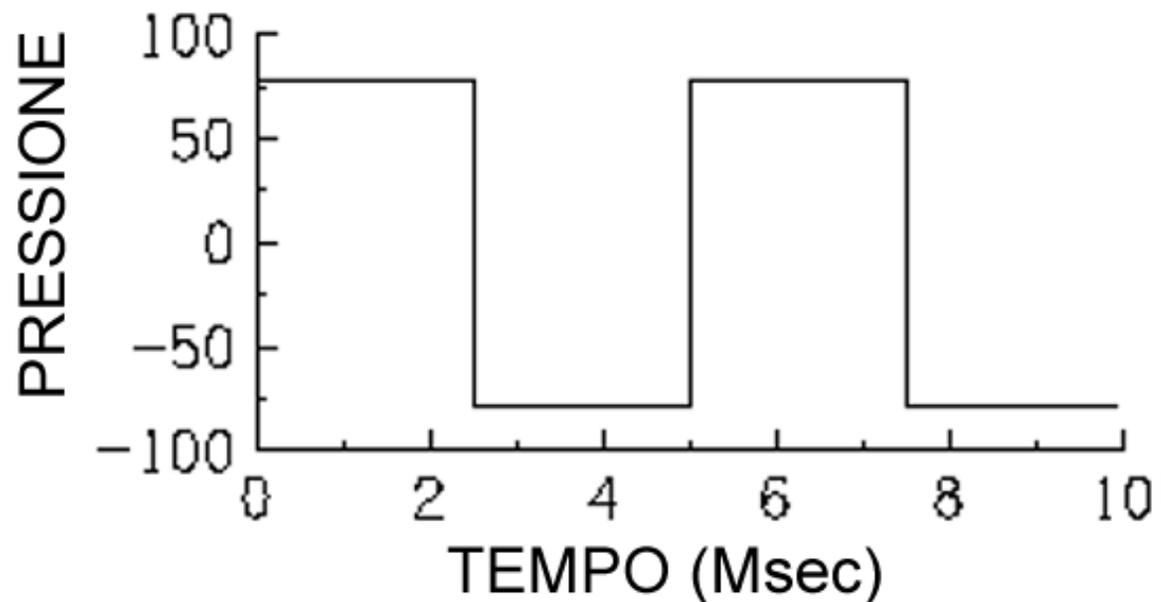
Scala lineare	X	Scala logaritmica	Y
1	1	1	10
2	2	2	100
3	3	3	1000
...
n	n	n	10^n

Proprietà fisiche e caratteristiche percettive 4

- **I suoni elementari hanno andamento sinusoidale, periodico, con estensione indefinita**
- Tuttavia, la maggior parte dei suoni presenti in natura non possono essere descritti così semplicemente essendo caratterizzati da forme d'onda diverse da quelle sinusoidali.
- Si può dimostrare che, fatte alcune ipotesi di regolarità sull'andamento della forma d'onda, **un generico suono complesso può essere descritto come una combinazione di suoni elementari.**

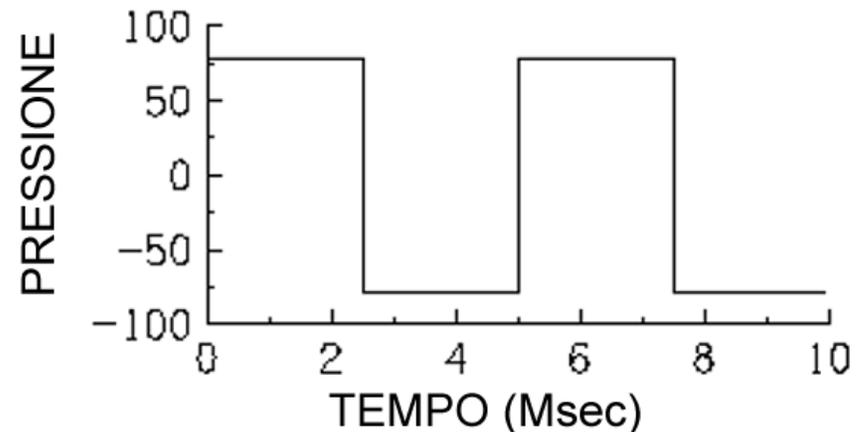
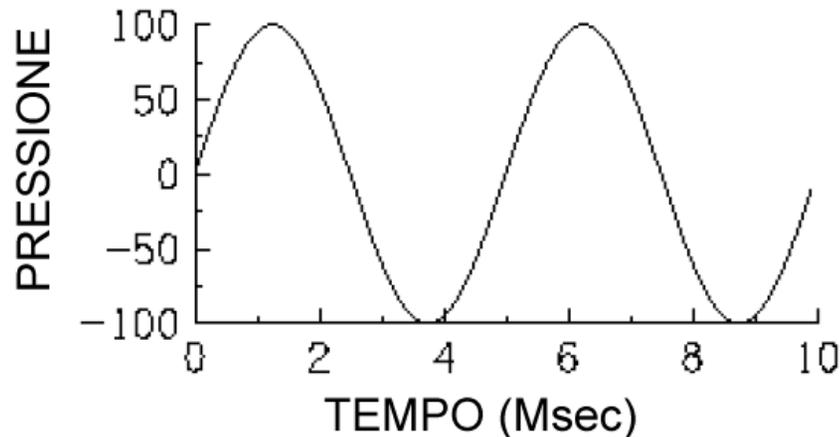
Proprietà fisiche e caratteristiche percettive 5

- Il suono corrispondente alla forma d'onda riportata in figura e' caratterizzato dall'alternarsi di intervalli temporali in cui la pressione assume valori costanti 'alti' e 'bassi'.



Proprietà fisiche e caratteristiche percettive 6

- Confronto tra due onde con la stessa frequenza: sebbene il tono delle due forme d'onda sia lo stesso, esse hanno un diverso timbro.



Proprietà fisiche e caratteristiche percettive 7

- Ogni suono può essere descritto come la somma di una serie di suoni elementari.
- L'onda quadra può essere ottenuta sommando onde sinusoidali di opportuna frequenza, ampiezza e fase
- Comprensibilmente, il primo e più rilevante contributo è dato dalla sinusoide di frequenza pari a quella dell'onda quadra: 200 Hz. Questa componente prende il nome di **FREQUENZA FONDAMENTALE** (F_0) ed è quella che determina il tono del suono percepito. Quindi la componente F_0 è quella che causa lo stesso tono percepito per la sinusoide a 200 Hz e per l'onda quadra a 200 Hz.
- Tuttavia, per costruire l'onda quadra sono necessarie anche altre componenti elementari di frequenza maggiore. Queste altre componenti vengono dette **ARMONICHE** ed hanno tutte frequenze multiple di F_0 , quindi 400 Hz, 600 Hz, 800 Hz e così via.

Proprietà fisiche e caratteristiche percettive 8

- L'onda quadra e' un caso particolare in cui tutte le armoniche pari sono nulle.

- Qui l'onda quadra e' data dalla somma delle componenti:

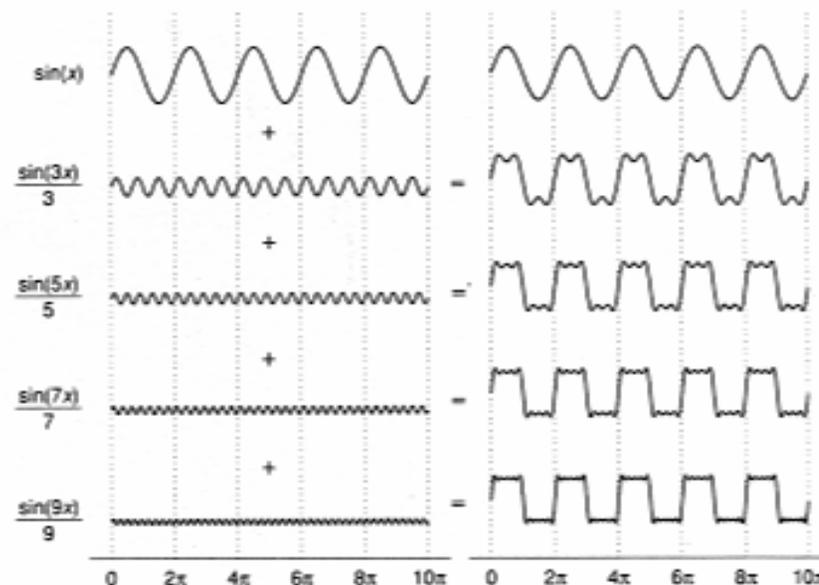
$$F_0, 3 * F_0, 5 * F_0 \text{ etc.}$$

vale a dire:

200Hz, 600Hz, 1KHz, 1.4KHz

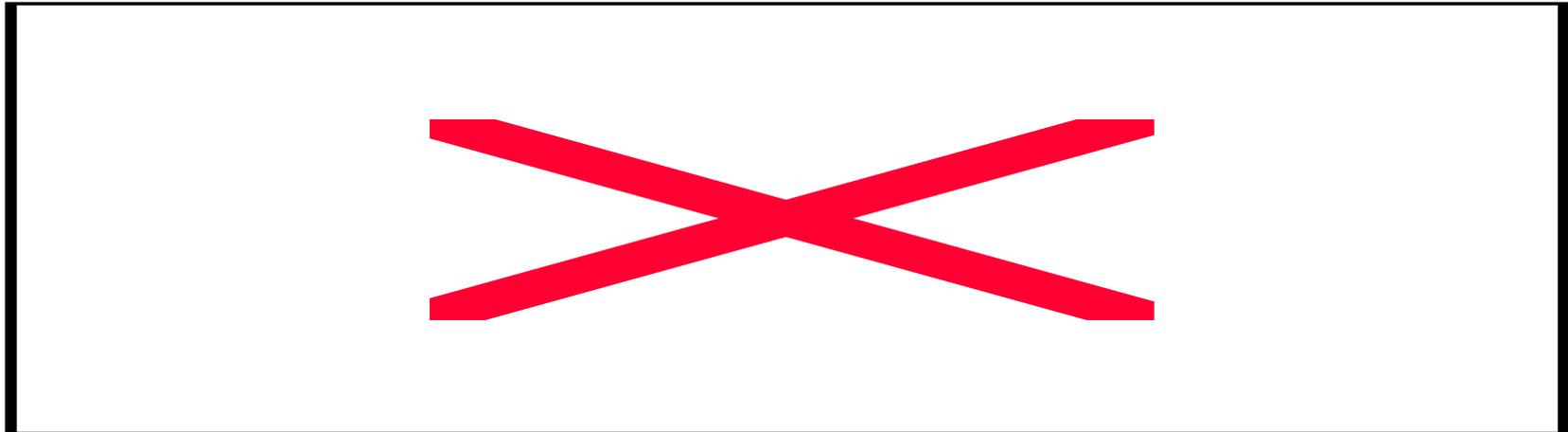
etc.

- Ognuna di queste componenti ha una diversa ampiezza rispetto a quella dell'armonica fondamentale



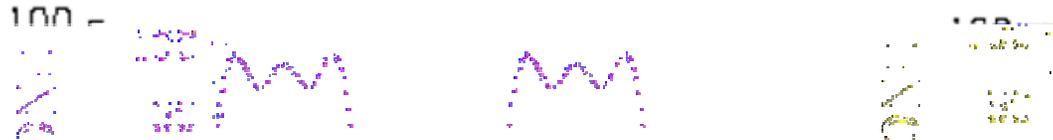
Proprietà fisiche e caratteristiche percettive 10

- L'onda ottenuta sommando i contributi della sinusoide a 200Hz con quella a 600Hz ad $1/3$ della sua ampiezza (a sx) e' la seguente (a dx):



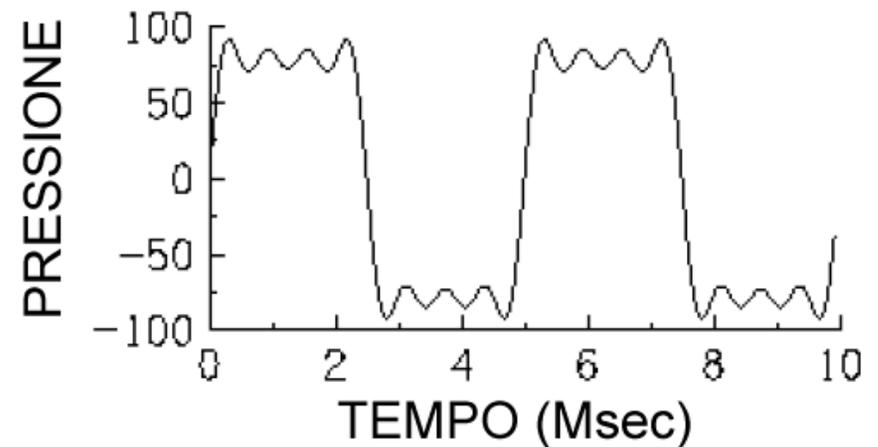
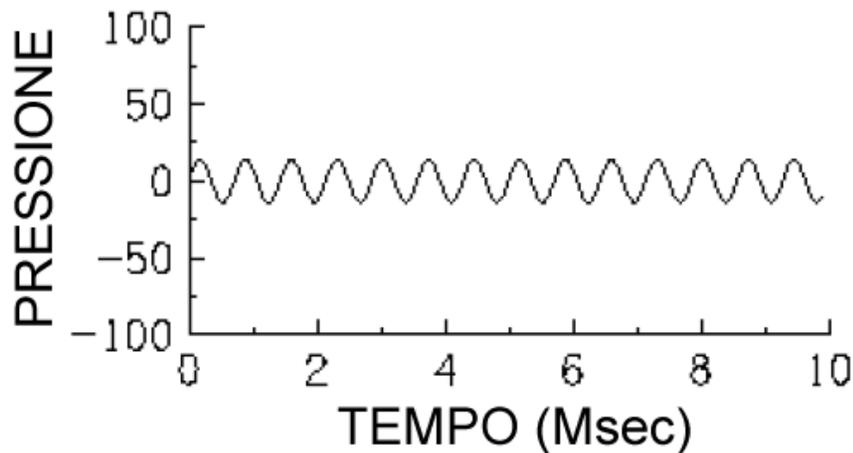
Proprietà fisiche e caratteristiche percettive 9

- continuando a sommare contributi di armoniche di frequenze via via crescenti, si arriva alla terza componente (la 5a armonica che ha 1000Hz) la cui ampiezza deve essere $1/5$ della fondamentale.



Proprietà fisiche e caratteristiche percettive 11

- Quindi abbiamo la 7a armonica (1.4KHz) con ampiezza $1/7$ della fondamentale

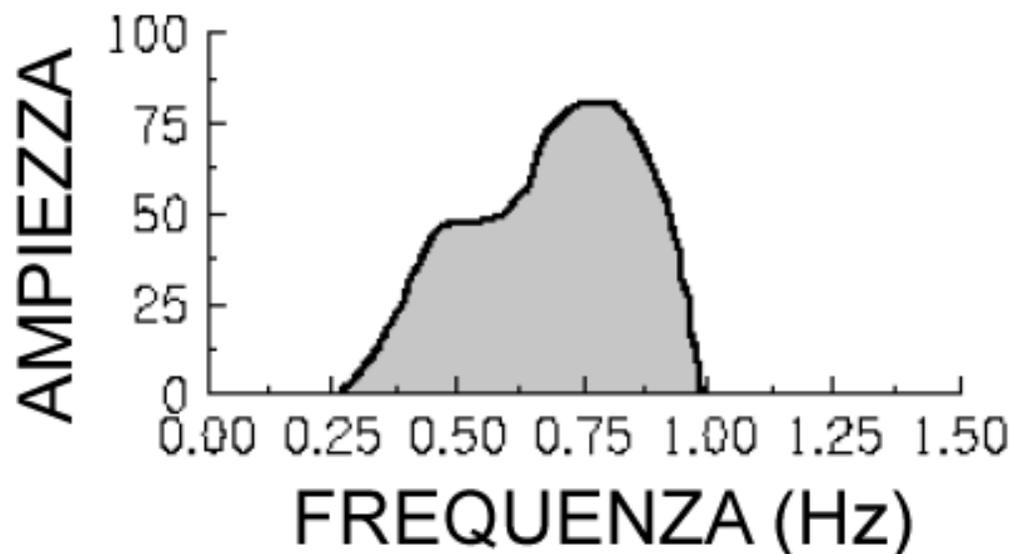


Rappresentazione tempo – frequenza (1)

- La presentazione sino ad ora effettuata e' basata su una descrizione dell'effetto sonoro in base a variazioni di pressione nel tempo.
- La rappresentazione nel dominio del tempo e' sicuramente importante ma per molti versi non rende esplicite alcune caratteristiche che sono invece di basilare importanza nella determinazione dell'effetto sonoro.
- Una di queste caratteristiche e' la frequenza, il cui valore e' di difficile derivazione da una rappresentazione della forma d'onda nel dominio del tempo.

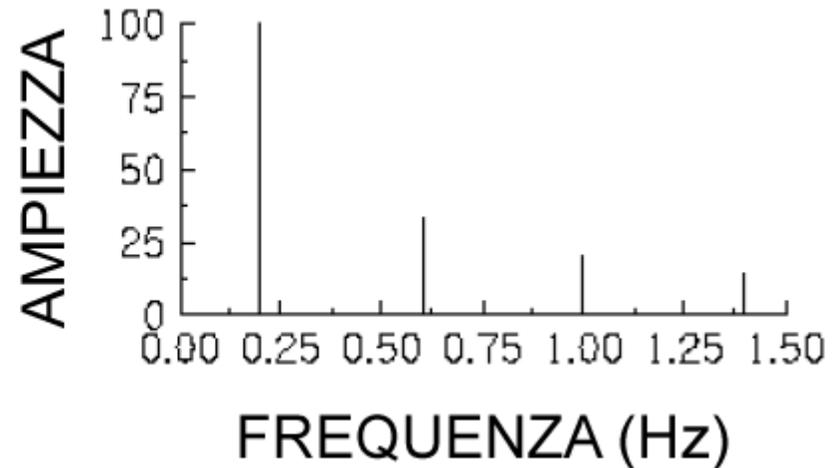
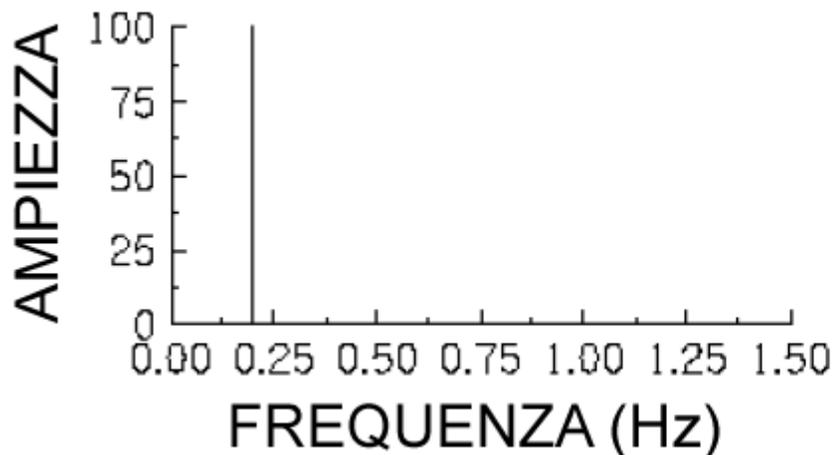
Rappresentazione tempo – frequenza (2)

- Per far fronte a queste difficoltà e' stata introdotta in una diversa rappresentazione del suono detta appunto rappresentazione nel dominio della frequenza.
- In questa rappresentazione, l'asse dei tempi viene sostituito dall'asse delle frequenze e la rappresentazione di un generico suono e' effettuata con un grafico che riporta i valori di ampiezza di ciascun contributo in frequenza.



Spettro Discreto e Spettro Continuo (1)

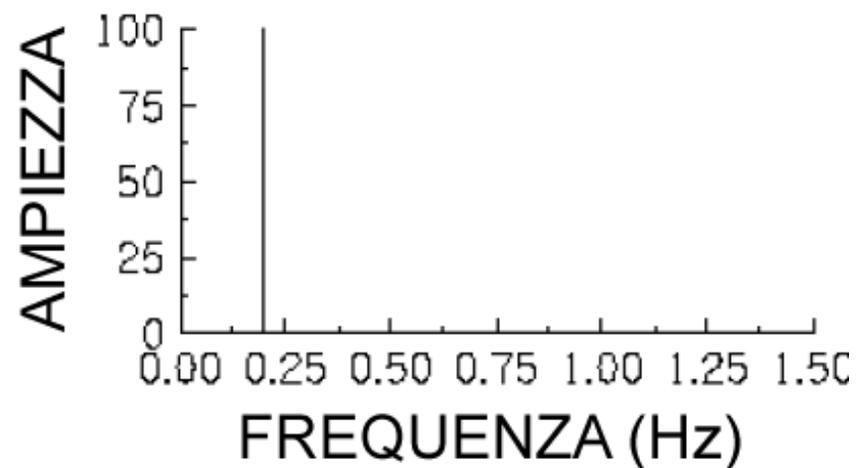
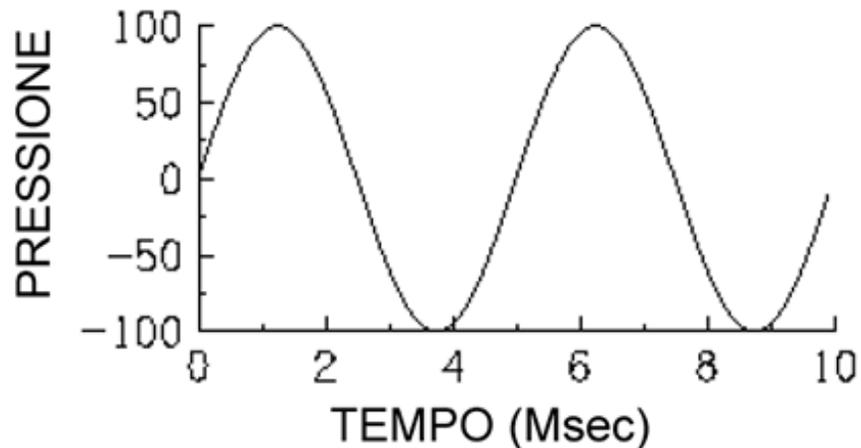
- Qui di seguito e' riportata la rappresentazione nel dominio della frequenza dell'ampiezza della forma d'onda sinusoidale e dell'onda quadra a 200 Hz



Si ottengono le linee spettrali che costituiscono uno spettro discreto nel dominio della frequenza

Spettro Discreto e Spettro Continuo (1)

- Lo spettro di un segnale complesso periodico è **discontinuo**: viene detto **spettro a righe** o **discreto**. La distribuzione e l'altezza delle righe spettrali è caratteristica del particolare segnale complesso periodico considerato. Lo spettro di un suono puro (cioè sinusoidale) per esempio, è costituito da una sola riga.



Spettro Discreto e Spettro Continuo (2)

- i suoni caratterizzati da uno spettro discreto sono pochi e **la maggior parte dei suoni presenti in natura presentano un andamento non periodico.**
- Dalla teoria dei segnali sappiamo che **un segnale perfettamente periodico deve avere una durata infinita**
- I suoni che normalmente si sentono hanno un inizio e una fine, vale a dire sono contenuti in un intervallo temporale finito

Spettro Discreto e Spettro Continuo (3)

- Esempio: rappresentazione in frequenza della forma d'onda quadra troncata al di fuori di un intervallo temporale finito.

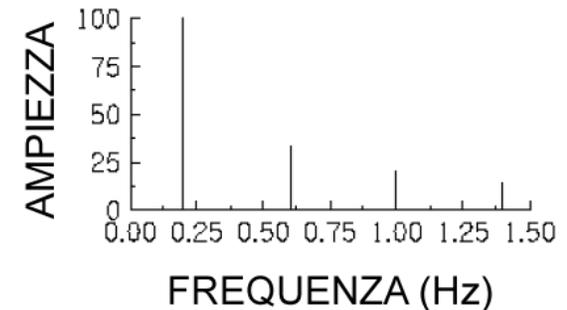


Spettro Discreto e Spettro Continuo (4)

- Osservazione: un'altra causa di scostamento rispetto ad un perfetto andamento periodico e' rappresentata dal fatto che la maggior parte dei suoni, anche all'interno dell'intervallo temporale in cui sono definiti, non hanno un andamento periodico ma **QUASIPERIODICO**, vale a dire presentano delle lievi differenze tra un periodo e l'altro.

Spettro Discreto e Spettro Continuo (5)

- La piu' importante differenza tra le due figure e' che nella prima vi sono delle linee armoniche, nella seconda si sono allargate in una forma a campana.
- Tali forme rappresentano sempre delle armoniche ma non associate a suoni puri. Rappresentano invece un suono associato ed una energia distribuita attorno a quella dell'armonica pura centrale.
- Spettri di questo tipo vengono chiamati **spettri armonici** invece che linee spettrali (o spettri a righe)



Spettro Discreto e Spettro Continuo (6)

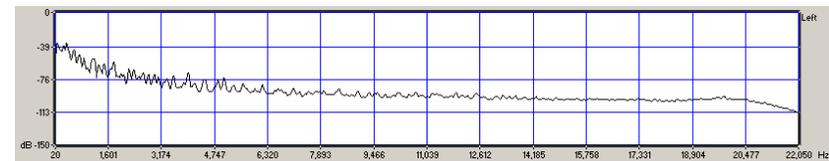
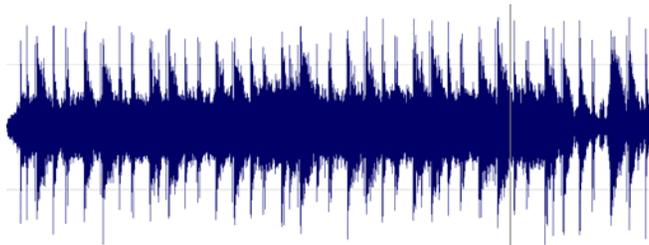
- La differenza tra le **righe** e le **campane** riflette una differenza fondamentale tra la rappresentazione nel dominio del tempo e quella nel dominio della frequenza.
- I suoni che si estendono per tempi molto lunghi e con andamenti fortemente periodici nel dominio del tempo sono caratterizzati da profili molto concentrati nel dominio della frequenza (una sinusoidale di durata infinita e' infatti caratterizzata da una unica linea spettrale).
- suoni definiti su intervalli temporali ristretti manifestano un profilo molto allargato nel dominio della frequenza.

Spettro Discreto e Spettro Continuo (7)

- La situazione estrema e' quella rappresentata da un suono puramente impulsivo: con ampiezza nulla ovunque a parte in un istante temporale. Un suono di questo tipo e' caratterizzato da uno spettro in frequenza che ha valore costante non nullo (spettro completamente piatto).
- Uno spettro di questo tipo non e' detto spettro armonico ma spettro **continuo**.
- Gli spettri continui sono associati a suoni con andamento temporale non periodico: **aperiodico**.

Spettro Discreto e Spettro Continuo (8)

- Un segnale complesso aperiodico è caratterizzato dall'aver uno **spettro continuo**, che comprende cioè un numero infinito di componenti distribuite con continuità sull'asse delle frequenze.
- Queste componenti, o parziali, non sono delle armoniche perché le loro rispettive frequenze non sono multipli interi di una stessa frequenza (manca quindi anche la fondamentale).
- La presenza dei contributi a frequenze elevate rende il suono più brillante, mentre la loro assenza lo rende più cupo.

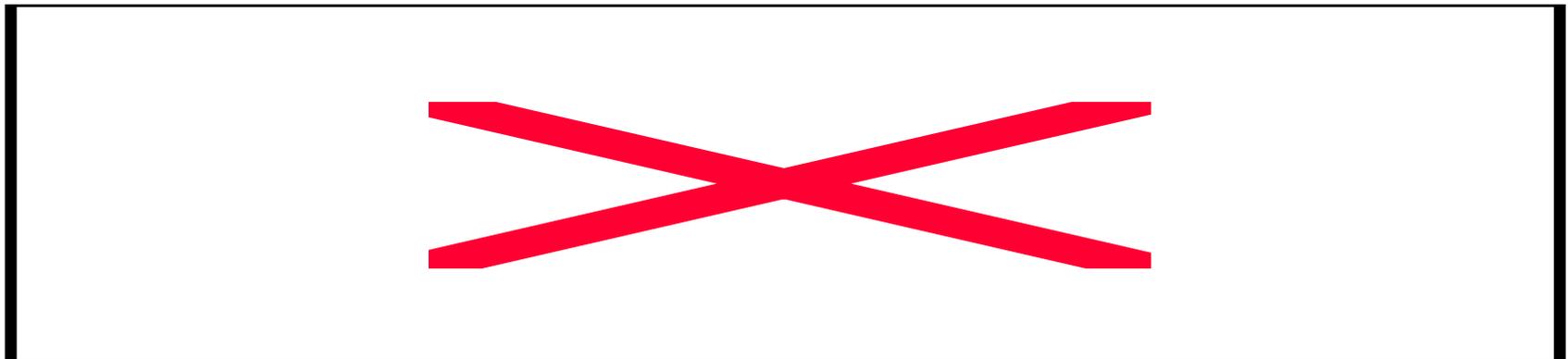


Il suono digitale (1)

- Il segnale analogico è quello in cui è continua la variazione nel tempo.
- Il segnale digitale è quello in cui la variazione nel tempo avviene invece in modo discreto.

Il suono digitale (2)

- I microfoni producono rappresentazioni analogiche del segnale audio. Questo e' infatti rappresentato da un valore di tensione il cui andamento nel tempo riflette le oscillazioni di pressione nell'aria

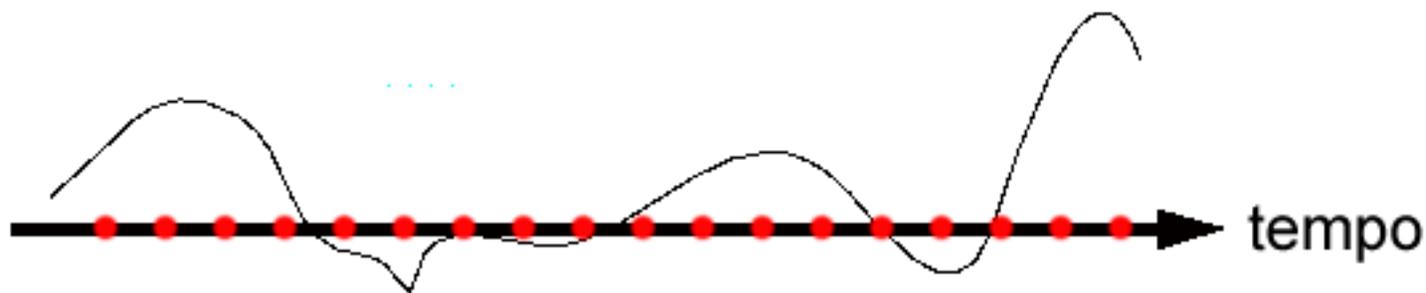


Il suono digitale (3)

- Per poter rappresentare il suono in un sistema digitale bisogna prima convertirlo in un flusso di numeri rappresentati in forma binaria.
- La conversione del suono da formato analogico a digitale avviene per mezzo di una scheda di acquisizione (o digitalizzazione) che campiona il valore della forma d'onda ad intervalli regolari.
- Inoltre, l'ampiezza di ogni campione, dovendo essere rappresentata digitalmente (cioè con una codifica binaria), non potrà assumere infiniti valori.
- La conversione **Analogico-Digitale** richiede pertanto un processo di discretizzazione sia nel tempo (**campionamento**) che in ampiezza (**quantizzazione**).

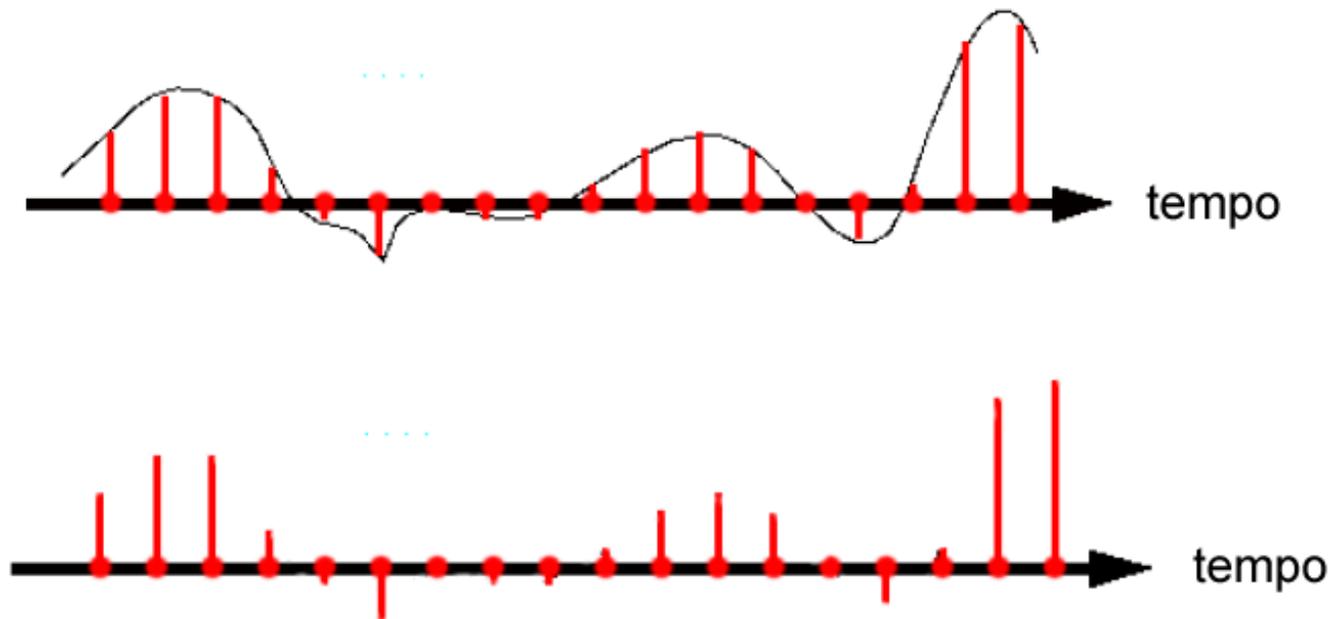
Campionamento

- Divisione dell'asse orizzontale (quello temporale) in unita' discrete. Il numero di campioni individuato in un secondo prende il nome di frequenza di campionamento.



Quantizzazione

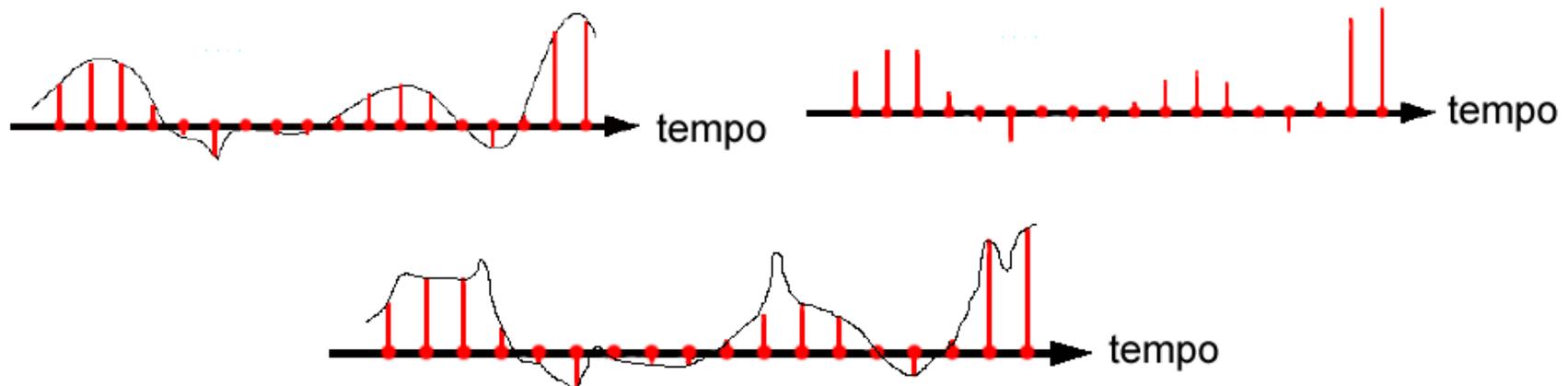
- divisione dell'asse verticale (quello delle ampiezze) in unita' di ampiezza costante



- Per la conversione Analogico-Digitale e' quindi necessaria la specifica di due parametri relativi a:
 - Quanto spesso campionare il segnale nel tempo (Frequenza di campionamento)
 - Con quanti valori rappresentare ogni campione (Precisione di quantizzazione)

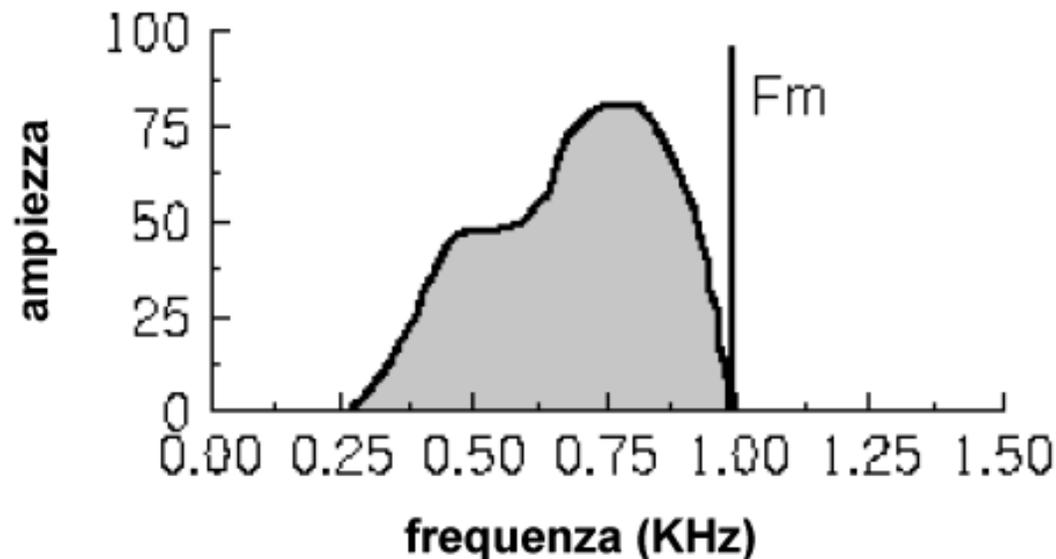
Ricostruzione del segnale

- Una volta estratti i campioni chi mi garantisce che questi individuino univocamente la forma d'onda di partenza?



Teorema di Nyquist

- Per avere una digitalizzazione senza perdita di informazione e' necessario campionare con una frequenza almeno il doppio della massima frequenza che compare nello spettro della forma d'onda da acquisire: La frequenza di campionamento di un segnale il cui spettro in frequenza abbia componenti non nulle sino a F_m dovrà'essere pari almeno a $2 \cdot F_m$.



- Sistema uditivo in grado di percepire suoni di frequenza compresa tra 20Hz e 22KHz.
- Range del parlato: 500Hz -2KHz
- un campionamento con frequenza maggiore di 44KHz dovrebbe garantire la perfetta riproducibilità del segnale audio.

Rapporto segnale rumore

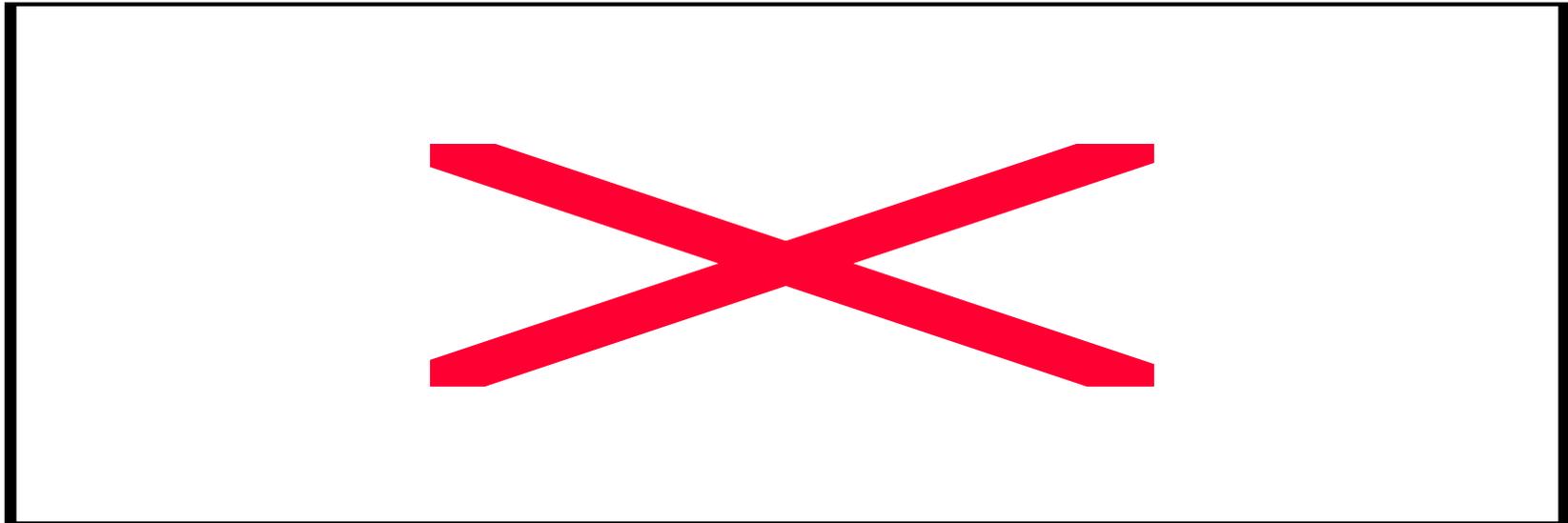
- Con qualsiasi rappresentazione analogica una parte del segnale e' dovuta al rumore.
- Il rapporto tra la potenza del segnale e quella del rumore e' chiamato **signal to noise ratio (SNR)**.
- SNR e' una misura della qualita' del segnale e si misura in decibel (dB):

$$SNR = 10 \log_{10} \left(\frac{\text{Signal Energy}}{\text{Noise Energy}} \right) = 10 \log_{10} \left(\frac{E_S}{E_N} \right)$$

Precisione di quantizzazione

- Tipicamente vengono usati 8 o 16 bit. Con 8 bit si hanno 256 distinti valori, con 16 bit si hanno 65536 distinti valori.
- Ogni bit aggiunge approssimativamente 6 dB di risoluzione, per cui 16 bit consentono di avere un rapporto segnale rumore $SNR = 96$ dB.
- Esempio: $V_{signal} = 2 \exp 16$, $V_{noise} = 1$. $\implies 20 \times \log (65536) = 96$ db

Audio Quality vs. Data Rate



Dimensione del file Audio

- La seguente formula da' una stima della dimensione del file rappresentante un minuto di traccia audio:

$$(\text{Frequenza in Hz}) * (\text{n. bit per campione}) / 8 = \text{Bytes/sec}$$

$$\text{Bytes/sec} * 60 = \text{Bytes/min}$$

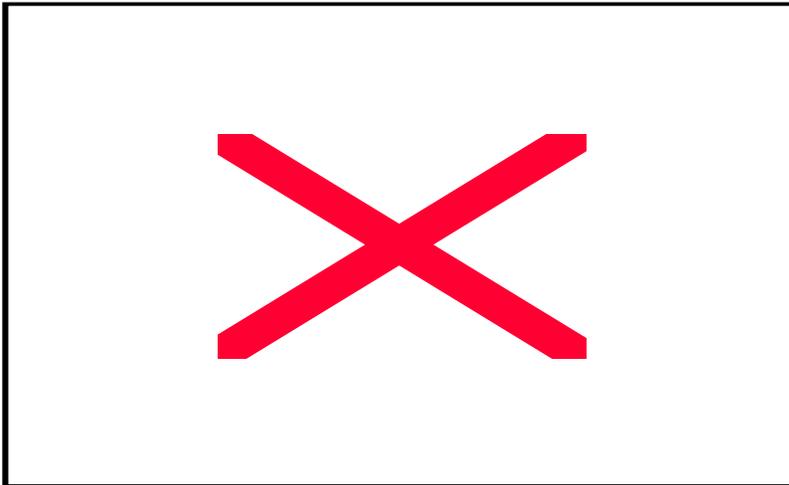
Dimensione del file Audio

- Di seguito sono riportate le dimensioni del file audio rappresentante una singola traccia di 5 minuti, per diversi valori di frequenza di campionamento e numero di bit per campione



Qualità audio

- La qualità audio del vinile è superiore a quella del CD e del DVD Audio, a causa della natura digitale della registrazione



	CD Audio	DVD Audio
Sampling Rate	44.1 kHz	192 kHz
Samples per second	44,100	192,000
Sampling Accuracy	16-bit	24-bit
Number of Possible Output Levels	65,536	16,777,216

Formati e supporti

- I formati dei file audio sono stati sviluppati per standardizzare la riproduzione e la distribuzione di dati audio nei sistemi digitali
- I parametri che determinano i dati audio sono tre:
 - **sampling rate**, misurato in campioni/sec (Hz), per canale;
 - lunghezza e tipo di codifica della parola binaria, ovvero il **numero di bit per campione**;
 - **numero di canali**;

Formati

- I formati si dividono in due tipi:
 - con intestazione (**header**), autodescriventi;
 - senza intestazione (**headerless** o raw);
- L'intestazione contiene:
 - definizione codifica usata per i dati audio;
 - descrizione brano
 - dati di copyright
- L'intestazione inizia spesso con una parola chiave, e prosegue poi con i dati della codifica



Formati audio

Estensione	Origine	Parametri	intestazione
au, snd_...	NeXT, Sun...	Data canali, codifica...	V...

Supporti

- I campioni del segnale digitale possono essere registrati su una moltitudine di supporti, con caratteristiche fisiche e logiche diverse:
 - **fisica**: supporti magnetici, ottici elettronici;
 - **logica**: ogni supporto è caratterizzato da uno standard proprio per la struttura dei dati

Supporti per l'audio digitale

- **CD-AUDIO:** $F_c=44,1$ KHz, quantizzazione 16 bit, segnale stereo, velocità di trasferimento minima:
 $44.100*2*16=1.411.200$ bit/sec;
- **CD-ROM:** applicazioni multimediali, 700 MB di dati, diversi standard (Libro Giallo, Rosso, etc.);

Supporti per l'audio digitale

- **DVD-ROM:** sola lettura, da 4,7 Gb a 17 Gb di capacità, lettori più lenti di quelli per CD ma i dati sono più vicini e quindi velocità di trasferimento è maggiore;
- **DVD-Video:** formato per i film in digitale, contiene sistema di protezione, formato video MPEG-2;
- **DVD scrivibili:** 5 formati diversi in funzione del supporto fisico e delle specifiche dei dati;
- **DVD-Audio:** competitore del SuperCD (Sony-Philips), audio di qualità multicanale con protezione, contenuti extra (immagini, testo, menu, video, etc.), connessione a Internet, Fc fino a 192Khz, quantizzazione fino a 24 bit;

Supporti per l'audio digitale

- **DAT:** nasce come formato audio per poi divenire nel 1998 un formato generale per l'archiviazione dei dati, nastro magnetico di 4 o 8 mm, formati variabili da 2 a 40 Gb;
- **Hard-Disk:** la memorizzazione dei dati avviene tramite polarizzazione di sostanze ferromagnetiche, le performance dipendono dall'interfaccia verso il processore (IDE, ATA, SCSI, USB esterne più lente);

Le connessioni digitali

- Le aziende che si occupano di audio digitale nel corso degli anni, hanno proposto standard diversi
- Esistono quindi numerosi formati audio digitali e vari differenti sistemi di trasmissione dell'audio digitale, con una gran numero di diverse sigle e diversi formati hardware per quanto riguarda prese, spinotti e cavi di trasmissione.
- Più formati convivono tuttora a causa della loro oggettiva utilità in particolari applicazioni, per cui sullo stesso calcolatore possono coesistere, oltre agli ingressi ed uscite analogici, più connessioni digitali contemporaneamente.

Le connessioni digitali

- Tra macchine digitali il trasferimento dei dati avviene senza uscire dal dominio digitale, ovvero senza conversioni successive D/A-A/D.
- In questo modo si evita l'effetto di degradazione del segnale dovuto alla doppia conversione

Interfacce audio digitali

- **AES3 (AES/EBU):**

- è il primo formato che definisce uno standard fra i vari costruttori di macchine audio a livello professionale. Stabilito dall'AES (Audio Engineering Society) verso la metà degli anni '80 e aggiornato nel 1992, permette la trasmissione di due canali di informazione audio digitale, contenenti dati audio e non audio.
- L'interfaccia è capace di trasmettere parole audio lunghe fino a 24 bit. Nella maggior parte dei casi le parole audio non superano i 20 bit e in tale situazione i rimanenti 4 bit destinati all'audio possono essere usati per informazioni ausiliarie.



Interfacce audio digitali

- **S/P-DIF (Sony/Philips Digital Interface Format):**

- Nelle applicazioni semiprofessionali e casalinghe è l'interfaccia digitale più comune
- prevede di funzionare con cavi coassiali tipo video utilizzando connettori phono su una distanza massima di 10 metri.
- è in grado di trasmettere il segnale di SCMS (Serial Copy Management System) in maniera tale da evitare, sulle apparecchiature casalinghe, la possibilità di copie digitali in serie di prodotti preregistrati (CD, etc.)



Interfacce audio digitali

- **T-DIF 1 (Tascam)**

- è il formato utilizzato per la registrazione di dati audio digitali dalle apparecchiature multitraccia a cassetta Tascam
- Il formato dei connettori utilizzati per la trasmissione è il T-DIF 1
- per la connessione viene impiegato un connettore D-sub a 25 poli che trasmette 8 canali di audio digitale simultaneamente.



Interfacce audio digitali

- **Toslink**

- connettori ottici, introdotti dalla Toshiba.
- Il formato trasmesso attraverso i connettori ottici è lo S/P-DIF, con un sistema di trasmissione totalmente differente, per cui la conversione o la trasmissione diretta tra connettori S/PDIF coassiali e ottici è impossibile;
- Il formato ottico viene anche utilizzato per la trasmissione del formato ADAT: 24 bit per 8 canali complessivi.
- Il formato ADAT è presente sui registratori multitraccia ADAT (Alesis Digital Audio Tape) che registra 8 canali audio digitali su una cassetta S-VHS. Il connettore ottico serve a trasmettere questi 8 canali ad un altro ADAT oppure ad una apparecchiatura di registrazione/riproduzione compatibile.



Interfacce audio digitali

- **MADI (Multichannel Audio Digital Interface) :**
 - È un'estensione del protocollo AES/EBU per connettere digitalmente apparecchiature multicanale
 - Utilizza fino a 56 canali audio in modo seriale lungo un singolo cavo terminato BNC a 75 Ohm (o alternativamente su fibra ottica), per distanze fino a 50 metri (fino a 2 Km su fibra ottica)
 - E' ideale quando si usino registratori digitali multitraccia e si vogliono connettere a mixer digitali rimanendo nel dominio digitale.



Il protocollo MIDI

- Musical Instrument Digital Interface
- Il MIDI è un linguaggio standard di comunicazione digitale utilizzato per scambiare informazioni tra strumenti musicali.
- l'elenco di istruzioni MIDI viene detto protocollo MIDI ed è stato introdotto ufficialmente nel 1983 da un consorzio di produttori con l'intento di permettere a macchine - anche se prodotte da diverse marche - di comunicare tra loro.
- per essere in grado di comunicare mediante il protocollo MIDI, un dispositivo deve essere dotato di un hardware specifico, detto interfaccia MIDI, che fornisce le connessioni necessarie. Vedremo più avanti quali sono queste connessioni e come funzionano.

Il protocollo MIDI 2

- Un'interfaccia MIDI ha due diversi componenti:
 - Hw: specifica il collegamento fisico tra gli strumenti musicali, specificata la porta MIDI, specifica il cavo MIDI e gestisce i segnali elettronici che vengono trasmessi tramite il cavo stesso.
 - Sw: specifica il formato dei dati che codifica l'informazione che viaggia attraverso l'hardware. Un dato in formato MIDI non codifica i singoli campioni, come fa invece il formato audio. La codifica contiene le specifiche degli strumenti, la definizione dell'inizio e della fine di una nota, la frequenza di base e il volume del suono.

Formato dei dati MIDI

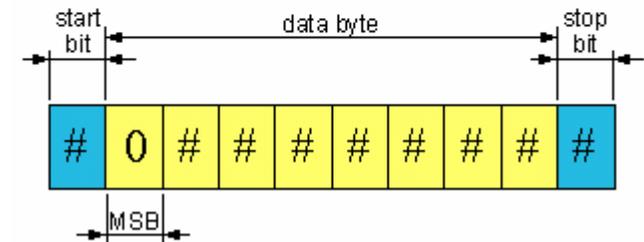
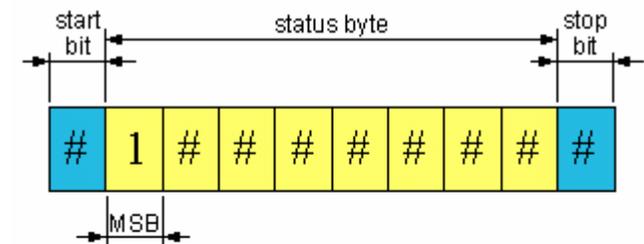
- Il formato dei dati MIDI è digitale: i dati vengono raggruppati in messaggi MIDI;
- Il protocollo MIDI non contiene informazioni audio: bisogna distinguere quindi il MIDI (.MID) dai formati audio (WAV, AIFF, etc.);

Formato dei dati MIDI 2

- Ogni messaggio MIDI comunica un evento musicale tra i diversi dispositivi collegati. Questi eventi sono tipicamente azioni che provoca un musicista mentre suona uno strumento musicale
 - esempio: quando un musicista preme il tasto di una tastiera elettronica l'interfaccia MIDI crea un messaggio MIDI codificando l'inizio della nota e l'intensità della pressione. Il messaggio viene trasmesso. Quando il tasto viene rilasciato, viene creato un nuovo messaggio MIDI, contenente il segnale corrispondente
- 10 minuti di musica sono circa 200KB di dati MIDI;

Formato dei dati MIDI 3

- I messaggi MIDI sono informazioni binarie che viaggiano in sequenza.
- Ogni "parola" è costituita da 10 bit, dei quali il primo e l'ultimo individuano l'inizio e la fine della parola, per cui il messaggio significativo della parola è contenuto negli 8 bit rimanenti
- Il primo bit viene detto start bit, i seguenti 8 data byte e l'ultimo bit stop bit.
- Un messaggio MIDI può essere costituito da una o da più parole successive contenenti 1 byte ciascuna.
- Il primo byte viene detto status byte e serve ad identificare il tipo di messaggio; il successivo o i successivi (non più di due, in ogni caso) contengono le informazioni significative.



Compressione Audio

- Le piu' moderne tecniche di compressione audio (AC3, MP3) si basano sulle caratteristiche dell'apparato uditivo umano. La principale caratteristica su cui si basano le migliori tecniche di compressione audio e' basata sul cosiddetto effetto di masking.
- Così' come nel dominio del visibile una forte sorgente luminosa tende ad abbagliare e nascondere la presenza di sorgenti luminose di minore intensita', la presenza di un forte suono ad certa frequenza tende a mascherare la presenza dei suoni a frequenze vicine.
- Queste tecniche di compressione si basano pertanto sulla suddivisione dello spettro audio in bande di frequenza di dimensione opportuna rispetto alla selettivita' del nostro sistema uditivo. Per ogni banda viene poi applicata una rappresentazione che tiene conto piu' che di tutte le componenti presenti nella banda, solo di quelle che sono effettivamente udibili (non mascherate).

Schemi di compressione semplici

- Gli schemi di compressione semplici son quelli con perdita di informazione che si basano su algoritmi non sofisticati
- Non raggiungono grandi compressioni, ma sono di rapida esecuzione e semplici da implementare.
- Esempi: standard μ -law e A-law, schema ADPCM

μ -law e A-law

- Schemi che realizzano le specifiche contenute nella raccomandazione G.711 rilasciata dal CCITT, comitato standard per le TLC;
- La codifica μ -law è utilizzata in Nord America e Giappone per i servizi di telefonia ISDN, A-law è usato invece in Europa e sul traffico internazionale ISDN;
- Frequenza di campionamento: 8 KHz;
- Quantizzazione logaritmica a 8 bit;
- Bitrate: 64 Kbps

Codifica ADPCM

- La codifica ADPCM (adaptive differential pulse code modulation) è un metodo comune di compressione che raggiunge un buon compromesso tra velocità di elaborazione, qualità del segnale decodificato e rapporto di compressione;
- Si basa su due concetti:
 - codifica delle differenze tra campioni (campioni adiacenti sono molto simili nel valore);
 - adattamento delle differenze allo specifico segnale di input
- invece di rappresentare ogni campione audio in modo indipendente, rappresenta la differenza tra un campione e un valore predetto a partire dal campione precedente e da un insieme di differenze quantizzate adattate sul segnale di input

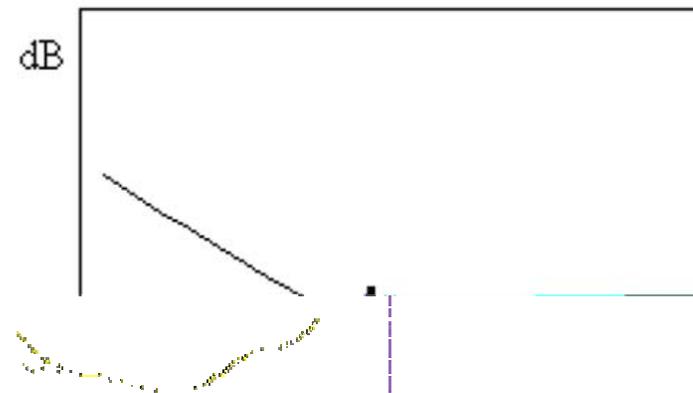
Schemi di compressione di tipo percettivo

- Gli schemi di compressione di questo tipo comprimono il segnale eliminando quelle parti che il nostro apparato uditivo non percepirebbe
- Sono codifiche di tipo lossy
- Tali schemi si basano sull'effetto di mascheramento

Effetto di mascheramento

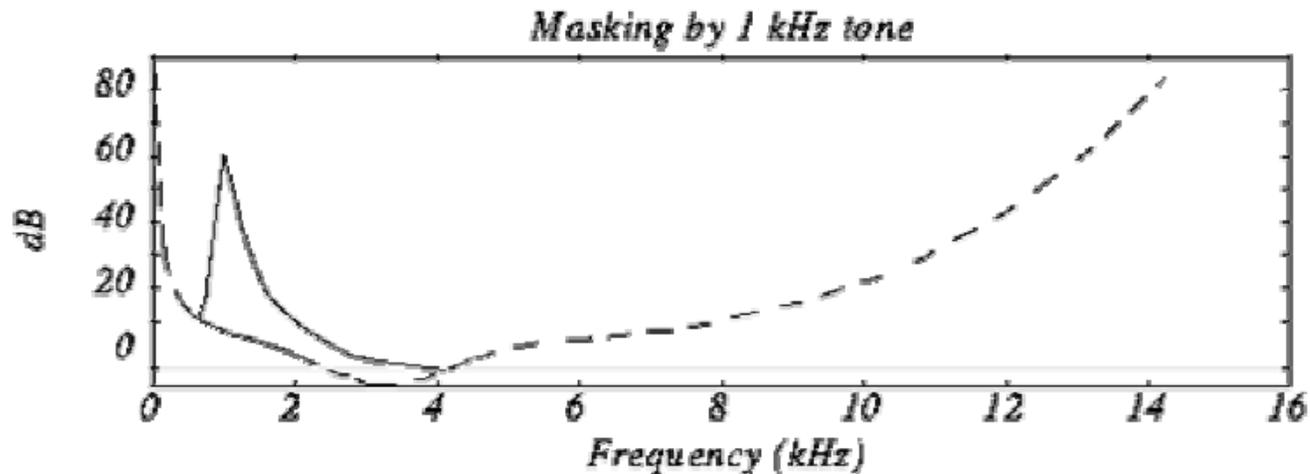
- Una forte componente di suono può rendere non udibili i suoni a frequenze vicine:

la presenza di una qualunque componente udibile in frequenza provoca una alterazione locale (e temporale) della soglia di udibilità attorno a quella frequenza.



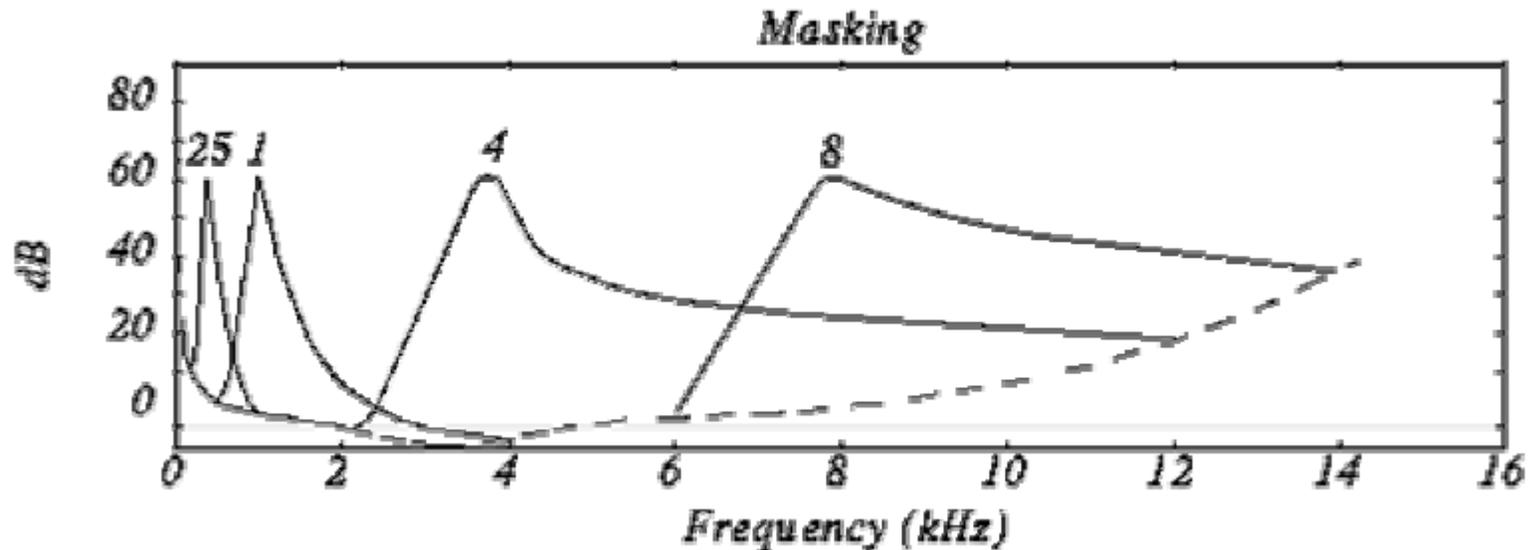
Effetto di mascheramento

- L'effetto di masking varia con la frequenza:
 - 100Hz per $f_0 = 20\text{Hz}$
 - 4KHz per $f_0 = 20\text{KHz}$.



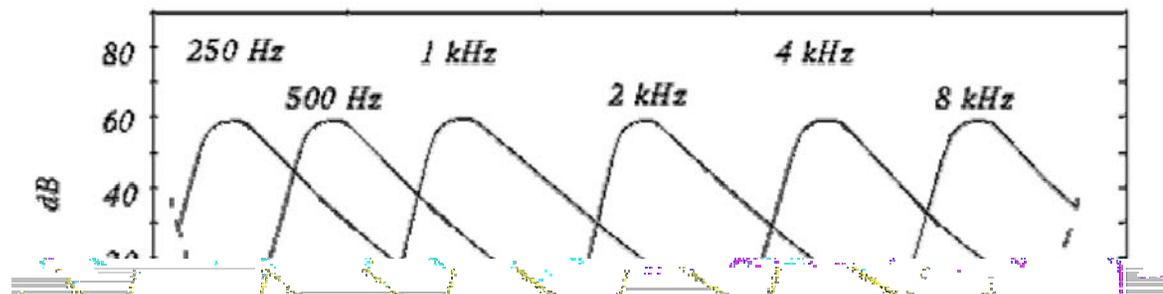
Effetto di mascheramento

- Circa 100Hz per frequenze di mascheramento < 500Hz, cresce per frequenze superiori.



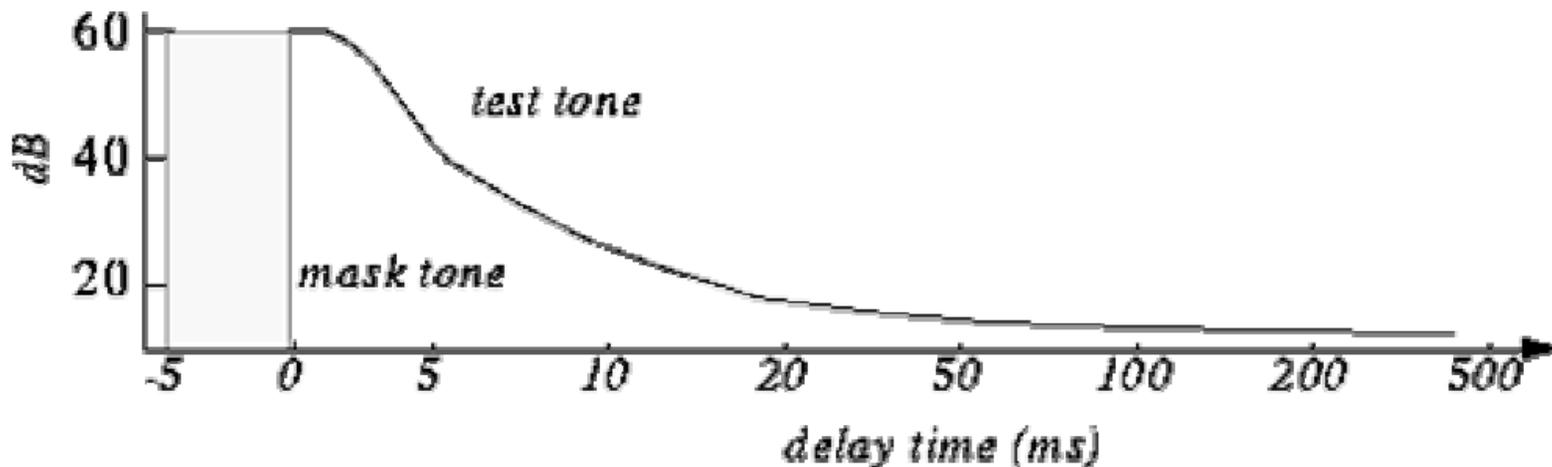
Banda critica

- Il range di frequenza per cui si manifesta l'effetto di mascheramento è detto banda critica.
- L'introduzione delle bande critiche permette di misurare la frequenza in un modo che sia uniforme dal punto di vista percettivo
- Nuova unità di frequenza:
- 1Bark = ampiezza della banda critica.
 - Per frequenze $< 500\text{Hz}$, $1 \text{ Bark} \sim \text{freq}/100$
 - per frequenze $> 500\text{Hz}$, $1\text{Bark} \sim 9 + 4\log(\text{freq}/1000)$



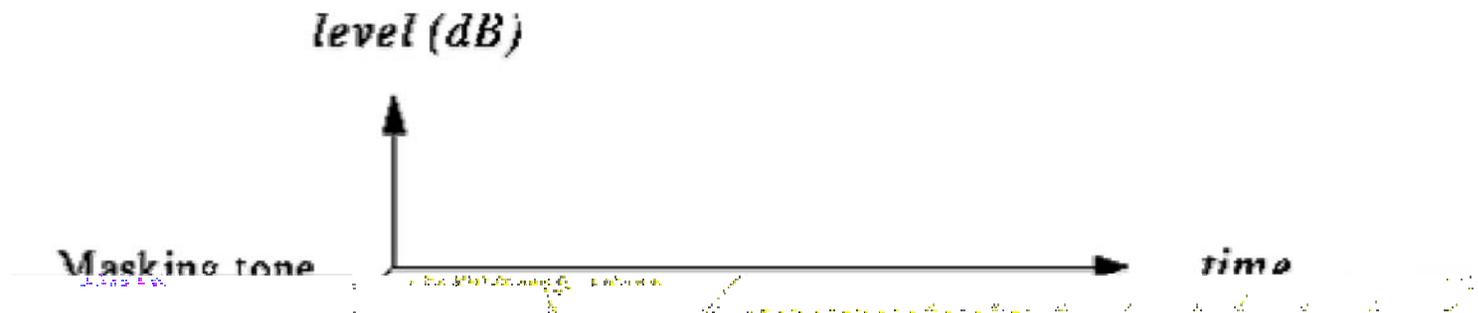
Mascheramento temporale

- Se ascoltiamo un suono di forte intensità, è necessario un certo tempo prima che un tono lieve possa essere percepito:
 - tono di mascheramento di 60dB a 1KHz;
 - tono di test a 1.1KHz a 40 dB;
- Ritardo di ascolto al variare dell'intensità del tono di test.



Effetto complessivo di mascheramento

- Variando la frequenza del tono di test, con tono di mascheramento di durata costante si ottiene il grafico seguente.



Banda critica

- Variazione della banda critica con la frequenza.



MPEG-1 audio

- Mpeg-1: 1.5 Mbits/sec per audio e video.
- Circa 1.2 Mbits/sec per la codifica video, 0.3 Mbits/sec per la codifica audio.
- Come confronto un CD audio non compresso ha un bit-rate:
 - $44100 \text{ campioni/sec} * 16\text{bits/campione} * 2 \text{ canali} >$
 - 1.4Mbits/sec.
- Fattore di compressione tra 2.7 e 24.
- Con fattore di compressione 6:1 (16 bit stereo campionati a 48KHz ridotti a 256 Kbits/sec) e condizioni di ascolto ottimali, ascoltatori esperti non sono in grado di distinguere tra clip audio codificati e originali.

MPEG-1 audio

- Mpeg-1: 1.5 Mbits/sec per audio e video.
- Circa 1.2 Mbits/sec per la codifica video, 0.3 Mbits/sec per la codifica audio.
- Come confronto un CD audio non compresso ha un bit-rate:
 - $44100 \text{ campioni/sec} * 16\text{bits/campione} * 2 \text{ canali} >$
 - 1.4Mbits/sec.
- Fattore di compressione tra 2.7 e 24.
- Con fattore di compressione 6:1 (16 bit stereo campionati a 48KHz ridotti a 256 Kbits/sec) e condizioni di ascolto ottimali, ascoltatori esperti non sono in grado di distinguere tra clip audio codificati e originali.

MPEG-1 audio

- Layer III dello standardMpeg-1: è il formato comunemente indicato come mp3.
- Prevede la compressione di uno stream con le seguenti frequenze di campionamento: 32KHz, 44.1KHz, 48KHz.
- Il formato audio di ingresso può essere in uno di quattro possibili modi:
 - Audio mono;
 - due canali per due traccemono;
 - due canali per una traccia stereo;
 - due canali per una tracciajoint-stereo (si avvantaggia della correlazione tra canali stereo).

MPEG-1 audio

- Sono previsti tre distinti livelli di compressione:
 - layerI: è il più semplice e maggiormente indicato qualora non sia richiesto uno stream di uscita inferiore a 128Kbit/s;
 - layerII: livello di complessità intermedio da impiegarsi qualora siano richiesti flussi di uscita di poco inferiore a 128Kbit/s. Le schede di acquisizione MPEG in real-time tipicamente supportano questo livello di compressione audio;
 - layerIII: è il più complesso e l'unico in grado di scendere sotto i 64Kbit/s mantenendo una elevata qualità audio.
- Per tutti e tre i livelli esistono chip che effettuano la decodifica in real-time.

Plugin

- Un plugin audio è un software che deve essere utilizzato all'interno di un'applicazione "host"
- Viene utilizzato per creare effetti sul suono

Plugin

- **AU (AudioUnit)**

- è la tecnologia per il controllo dei plugin introdotta da Apple per Mac OS X.
- vengono utilizzati, ad esempio, dal sequencer Logic Pro di Apple



Plugin

- **DirectX**

- è l'insieme dei driver distribuiti da Microsoft per i sistemi operativi Windows, per la gestione di grafica, audio e animazioni.
- lavorano in real time
- lavorano in real time (ossia, possono processare o alterare il suono senza creare un nuovo file).



Plugin

- **TDM (Time Domain Multiplex)**
 - è un sistema per la gestione dei plugin creato da Digidesign per i sistemi Pro Tools
 - richiede l'uso di una scheda audio della Digidesign per poter funzionare
 - lavora con 24 bit e 256 canali, integrando il mixing ed il processing in real time dei segnali digitali
 - I plugin TDM possono anche essere utilizzati su programmi host come Logic Audio e Digital Performer, ma sempre in combinazione con l'uso di una interfaccia Digidesign.

Plugin

- **HTDM (Host TDM)**
 - si riferisce a plugin che fanno tutto il lavoro di elaborazione dell'audio sulla CPU dell'host, invece che sui chip DSP disponibili sull'hardware TDM
 - consente una maggiore potenza e flessibilità per l'utilizzo di synth e campionatori integrati direttamente in Pro Tools.
 - In pratica questi plugin vengono utilizzati come normali plugin TDM, ma includendo i benefici di utilizzare l'automazione di Pro Tools



Plugin

- **MAS (MOTU Audio System)**
 - E' una tecnologia introdotta da MOTU (MArk Of The Unicorn) per la gestione dei plugin in real-time sui propri software, come Digital Performer
 - sono completamente automatizzabili e non richiedono chip DSP esterni
 - solo per MacOS

Plugin

- AudioSuite
 - sono stati sviluppati molto tempo fa, non possono essere utilizzati in tempo reale
 - gli effetti devono essere applicati quando il software non è in riproduzione e solo in seguito ascoltare il risultato;
 - Il vantaggio è che sono in grado di effettuare operazioni molto complesse senza sovraccaricare il sistema

Plugin

- **RTAS (Real Time AudioSuite)**

- aggiunge ai plugin Audio Suite la possibilità di utilizzare gli effetti in tempo reale (solo Mac)
- offre performance analoghe a quelle del sistema TDM
- ha superato AudioSuite

Plugin

- **VST (Virtual Studio Technology)**
 - E' una tecnologia real time introdotta nel 1996 da Steinberg per consentire di gestire i plugin all'interno del software Cubase
 - Sia Mac che Windows



Media Integration and Communication Center

Università degli Studi di Firenze

contatti

Gianpaolo D'Amico

sito web: <http://morpheus.micc.unifi.it/damico/ppm>

indirizzo e-mail: damico@dsi.unifi.it